

# Vinay Pratap Singh

Software Engineering Lead

## PROFILE

**Around 7.5 years** of experience in Information Technology industry with 6.5 years working on **Big Data**. Previously worked in Accenture, now working in I2E Consulting.

Domains – **Pharma**

**Retail**

**Financial Services**

Worked on **SDLC** Waterfall model and **Agile** methodology.

Briefly worked on various phases of Project Life Cycle like Requirement Gathering, Analysis, Design, Build, Testing, Implementation, Post Implementation Support.



Noida, U.P.



[vinayps91@gmail.com](mailto:vinayps91@gmail.com)

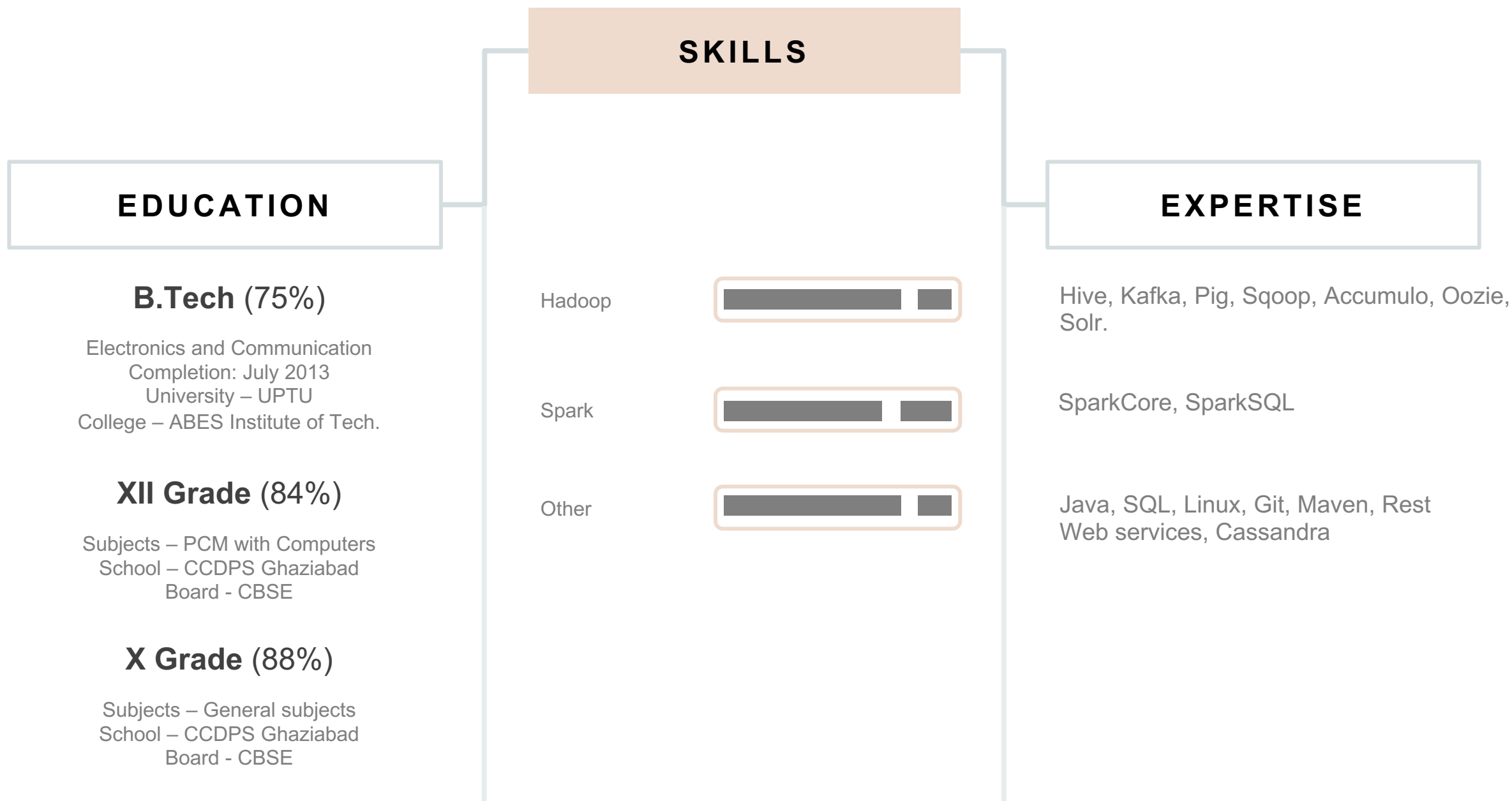


-



+91-9987840800





## EDUCATION

### B.Tech (75%)

Electronics and Communication  
Completion: July 2013  
University – UPTU  
College – ABES Institute of Tech.

### XII Grade (84%)

Subjects – PCM with Computers  
School – CCDPS Ghaziabad  
Board - CBSE

### X Grade (88%)

Subjects – General subjects  
School – CCDPS Ghaziabad  
Board - CBSE

## SKILLS

Hadoop



Spark



Other



## EXPERTISE

Hive, Kafka, Pig, Sqoop, Accumulo, Oozie, Solr.

SparkCore, SparkSQL

Java, SQL, Linux, Git, Maven, Rest Web services, Cassandra

## EXPERIENCE

### Big Data Head

- Currently handling the big data vertical of the company which includes pre-sales, current projects and maintenance of projects.
- Heading the big data applications for the company and managing different teams collaboratively to achieve high efficiency solutions for clients.
- Mostly involved in the beginning of the application from scratch and then my team takes care from there.
- Currently started working on building an application which integrates all the data from source systems using Hadoop.
- All data is ingested from other systems into **Hive** tables using **Sqoop**.
- All data transformations and processing happens in **Spark** on Hive tables including lot of joins on different entities partitioned by date.
- This data will later be used to generate reports and analytics purposes.

#### FRANCE PHARMA COMPANY

Aug 2020 – Jan 2021

I2E Consulting

Technologies – Spark with  
Scala, Hive, Sqoop

### Product Lead

- Handling a very big product called 'Cortellis Data Fusion' which was originally developed by Thomson Reuters, now we are developing features in it under a major US Pharma client.
- The product fetches data from all data sources of the company and make relations between them in the form of entities. Finally when the users search anything, it gives a final analysis of all the relations that item has from the information it gathered from all the data sources.
- The technologies used are Groovy with Grails, Java, Solr, Accumulo, Hadoop, PostGres, ExtJS, Gradle etc.

#### US PHARMA COMPANY

April 2018 – Current

I2E Consulting

Technologies - Groovy with  
Grails, Java, Solr, Accumulo,  
Hadoop, PostGres, Redis.

## EXPERIENCE

### Senior Software Engineer

- The project was a POC project. It was related to extract data from hundreds and thousands of big excels created by scientists and process that data according to business and save to AWS.
- Used **Java with Spark** to extract data from excels and put in **JavaRDDs** to process the data.
- Used **dataframes** to analyse the data in structured way.
- Used **AWS** to store the resulting output.

#### US PHARMA COMPANY

Feb 2018 to Mar 2018  
I2E Consulting

Technologies – Spark, Java,  
AWS Cloud.

### Senior Software Engineer

- The project was to develop middle-tier **Java Web Services** that can query **elastic search** index for real time searching.
- Used Java Rest Web Service to **add, search, delete and update the indexes** in elastic search.
- **Installed and Setup Elastic Search** from scratch and **setup indexes and mappings** for operations.

#### US PHARMA COMPANY

Dec 2017 to Jan 2018  
I2E Consulting

Technologies – Java, Elastic  
Search, REST Api

## EXPERIENCE

### Software Engineer

The project was to ingest data from around 30 applications into our application and export the data to Teradata application for reporting purposes.

- Used **Hive, Java & Pig** for end-to-end development of different data topics into different layers of application.
- Used **Pig** to process the clickstream logs data.
- **Design workflows** for java, hive and pig scripts in oozie.
- Use **Falcon feeds & processes** for triggering workflows on basis of data availability.
- Writing different **UDFs** for adding extra fields as surrogate keys in the data to uniquely identify rows, etc and used **Cassandra** for mapping of these transactions.
- Worked on **Avro** and **Json** data formats.
- **Installed & configured Kerberos** for HDFS taking the security to another level.
- **Installation & configured Encrypted** HDFS cluster.

#### US RETAIL COMPANY

Jan 2016 to Nov 2017

Accenture

Technologies – Hive, Pig,  
Oozie, Java, Cassandra,  
Kerberos, Rangers, Kafka  
#Hortonworks

### Software Engineer

- **POC Project** for analyzing **web logs** using Pig.
- Used different **UDFs** for concatenating, trimming, summing, counting, etc of data according to business logic customizations.
- Made scripts for analyzing data for various items in store.
- Capture user id and their interest item fields from the log data using **regular expressions**.

#### US INSURANCE COMPANY

Nov 2015

Accenture

Technologies – Pig, Java

## EXPERIENCE

### Software Engineer

The project was for ETL the records of insurance transactions from various systems daily.

- My role was to identify and automate technical processes.
- Within 2 months, Analyzed the pain point for the client and gave an idea for **Automation** and also implemented it single-handedly.
- The Automation Application written in **Java** extracts all the xml files from a big file and then extracts the values of each xml, and compares and validates with the source file and writes its decisions into a new excel file with formatting also automated.

#### US INSURANCE COMPANY

July 2015 to Nov 2015

Accenture

Technologies – Java

### Associate Software Engineer

The aim of the project was to remediate the customers who do not get bonus in any case but are eligible for it. So analysis is done on their transactions and status and their bonus is calculated to award them bonus manually.

- Having experience in Importing and Exporting data to HDFS using **SQOOP**.
- **Installation and running** of different components of Hadoop.
- Involved in creating hive tables with **partitions and bucketing**.
- Capable of processing large data that is **Data Extraction, Transformation and Data Loading** in Data Warehouse environment.
- Executed queries using **Pig** to analyze data to solve problems.
- Hands on experience in **joining** tables, **grouping** fields, **filtering** records.
- Developed **Pig & Hive UDF's** for gaining more control over the process of data.
- Worked on **Analytical Functions** like lead, lag, rank etc.
- Good knowledge of MapR Framework & HDFS Architecture.

#### US CREDIT CARD COMPANY

Jan 2014 to June 2015

Accenture

Technologies – Sqoop, Hive,  
Pig

## PERSONALITY

### AWARDS/EXTRA

Employee of the Year Award – I2E

Ace Award - Accenture

All India AMCAT Score – 96% in QA,  
99% in Logical Ability

Quick learner, Amicable, Self-motivator.

Enthusiastic to work in a team.

Quality to embrace change with agility and  
adaptability

Motivate group of the people to achieve  
organizational objectives.

### INTERESTS

Swimming

Playing Table-Tennis

Playing Football

Treking

Horse Riding